

RAW SEQUENCE LISTING

DATE: 10/19/2001

PATENT APPLICATION: US/09/921,650A

TIME: 12:07:21

Input Set : N:\Crif3\RULE60\09921650A.txt

Output Set: N:\CRF3\10192001\I921650A.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Bujard, Hermann
6 Gossen, Manfred

8 (ii) TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
9 Inhibitor Fusion Proteins

11 (iii) NUMBER OF SEQUENCES: 37

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: LAHIVE & COCKFIELD
15 (B) STREET: 28 State Street
16 (C) CITY: Boston
17 (D) STATE: Massachusetts
18 (E) COUNTRY: USA
19 (F) ZIP: 02109-1875

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk
23 (B) COMPUTER: IBM PC compatible
24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
25 (D) SOFTWARE: ASCII Text

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/921,650A
C--> 29 (B) FILING DATE: 03-Aug-2001
70 (C) CLASSIFICATION:

C--> 67 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: 09/489,777
34 (B) FILING DATE: 2000-01-24
38 (A) APPLICATION NUMBER: US 08/485,978
39 (B) FILING DATE: 07-JUN-1995
43 (A) APPLICATION NUMBER: US 08/383,754
44 (B) FILING DATE: 03-FEB-1995
48 (A) APPLICATION NUMBER: US 08/275,876
49 (B) FILING DATE: 15-JULY-1994
53 (A) APPLICATION NUMBER: US 08/270,637
54 (B) FILING DATE: 01-JULY-1994
58 (A) APPLICATION NUMBER: US 08/260,452
59 (B) FILING DATE: 14-JUNE-1994
63 (A) APPLICATION NUMBER: US 08/076,327
64 (B) FILING DATE: 14-JUNE-1993
68 (A) APPLICATION NUMBER: US 08/076,726
69 (B) FILING DATE: 14-JUNE-1993

72 (viii) ATTORNEY/AGENT INFORMATION:

73 (A) NAME: DeConti, Giulio A. Jr.
74 (B) REGISTRATION NUMBER: 31,503
75 (C) REFERENCE/DOCKET NUMBER: BBI-009C6CNDV

77 (ix) TELECOMMUNICATION INFORMATION:

78 (A) TELEPHONE: (617)227-7400
79 (B) TELEFAX: (617)742-4214

ENTERED

RAW SEQUENCE LISTING

DATE: 10/19/2001

PATENT APPLICATION: US/09/921,650A

TIME: 12:07:21

Input Set : N:\Crf3\RULE60\09921650A.txt

Output Set: N:\CRF3\10192001\I921650A.raw

81 (2) INFORMATION FOR SEQ ID NO: 1:
83 (i) SEQUENCE CHARACTERISTICS:
84 (A) LENGTH: 1008 base pairs
85 (B) TYPE: nucleic acid
86 (C) STRANDEDNESS: double
87 (D) TOPOLOGY: linear
W--> 89 (ii) MOLECULE TYPE: DNA
91 (ix) FEATURE:
92 (A) NAME/KEY: exon
93 (B) LOCATION: 1..1008
95 (ix) FEATURE:
96 (A) NAME/KEY: mRNA
97 (B) LOCATION: 1..1008
99 (ix) FEATURE:
100 (A) NAME/KEY: misc. binding
101 (B) LOCATION: 1..207
103 (ix) FEATURE:
104 (A) NAME/KEY: misc. binding
105 (B) LOCATION: 208..335
107 (ix) FEATURE:
108 (A) NAME/KEY: CDS
109 (B) LOCATION: 1..1005
111 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
113 ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG 48
114 Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
115 1 5 10 15
117 CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG 96
118 Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
119 20 25 30
121 AAG CTA GGT GTA GAG CAG CCT ACA CTG TAT TGG CAT GTA AAA AAT AAG 144
122 Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
123 35 40 45
125 CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT 192
126 Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
127 50 55 60
129 ACT CAC TTT TGC CCT TTA AAA GGG GAA AGC TGG CAA GAT TTT TTA CGC 240
130 Thr His Phe Cys Pro Leu Lys Gly Glu Ser Trp Gln Asp Phe Leu Arg
131 65 70 75 80
133 AAT AAG GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC AAT GGA 288
134 Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asn Gly
135 85 90 95
137 GCA AAA GTA CAT TCA GAT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT 336
138 Ala Lys Val His Ser Asp Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
139 100 105 110
141 CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG 384
142 Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
143 115 120 125
145 AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC 432
146 Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/921,650A

DATE: 10/19/2001

TIME: 12:07:21

Input Set : N:\Crf3\RULE60\09921650A.txt

Output Set: N:\CRF3\10192001\I921650A.raw

```

147      130      135      140
149 GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA      480
150 Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
151 145      150      155      160
153 CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA      528
154 Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu
155      165      170      175
157 TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG      576
158 Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
159      180      185      190
161 ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCC GCG      624
162 Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala
163      195      200      205
165 TAC AGC CGC GCG CGT ACG AAA AAC AAT TAC GGG TCT ACC ATC GAG GGC      672
166 Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly
167      210      215      220
169 CTG CTC GAT CTC CCG GAC GAC GAC GCC CCC GAA GAG GCG GGG CTG GCG      720
170 Leu Leu Asp Leu Pro Asp Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala
171 225      230      235      240
173 GCT CCG CGC CTG TCC TTT CTC CCC GCG GGA CAC ACG CGC AGA CTG TCG      768
174 Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser
175      245      250      255
177 ACG GCC CCC CCG ACC GAT GTC AGC CTG GGG GAC GAG CTC CAC TTA GAC      816
178 Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp
179      260      265      270
181 GGC GAG GAC GTG GCG ATG GCG CAT GCC GAC GCG CTA GAC GAT TTC GAT      864
182 Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp
183      275      280      285
185 CTG GAC ATG TTG GGG GAC GGG GAT TCC CCG GGT CCG GGA TTT ACC CCC      912
186 Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
187      290      295      300
189 CAC GAC TCC GCC CCC TAC GGC GCT CTG GAT ATG GCC GAC TTC GAG TTT      960
190 His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
191 305      310      315      320
193 GAG CAG ATG TTT ACC GAT CCC CTT GGA ATT GAC GAG TAC GGT GGG TAG      1008
194 Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly
195      325      330      335
197 (2) INFORMATION FOR SEQ ID NO: 2:
199 (i) SEQUENCE CHARACTERISTICS:
200 (A) LENGTH: 335 amino acids
201 (B) TYPE: amino acid
202 (D) TOPOLOGY: linear
204 (ii) MOLECULE TYPE: protein
206 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
208 Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
209 1      5      10      15
211 Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
212      20      25      30
214 Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/921,650A

DATE: 10/19/2001

TIME: 12:07:21

Input Set : N:\Crf3\RULE60\09921650A.txt

Output Set: N:\CRF3\10192001\I921650A.raw

```
215          35          40          45
217 Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
218          50          55          60
220 Thr His Phe Cys Pro Leu Lys Gly Glu Ser Trp Gln Asp Phe Leu Arg
221 65          70          75          80
223 Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asn Gly
224          85          90          95
226 Ala Lys Val His Ser Asp Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
227          100          105          110
229 Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
230          115          120          125
232 Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys
233          130          135          140
235 Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
236 145          150          155          160
238 Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu
239          165          170          175
241 Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
242          180          185          190
244 Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala
245          195          200          205
247 Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly
248          210          215          220
250 Leu Leu Asp Leu Pro Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala
251 225          230          235          240
253 Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser
254          245          250          255
256 Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp
257          260          265          270
259 Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp
260          275          280          285
262 Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
263          290          295          300
265 His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
266 305          310          315          320
268 Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly
269          325          330          335
```

272 (2) INFORMATION FOR SEQ ID NO: 3:

274 (i) SEQUENCE CHARACTERISTICS:

275 (A) LENGTH: 33 base pairs

276 (B) TYPE: nucleic acid

277 (C) STRANDEDNESS: double

278 (D) TOPOLOGY: linear

W--> 280 (ii) MOLECULE TYPE: DNA

283 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

285 GAC GCG CTA GAC GAT TTC GAT CTG GAC ATG TTG 33

286 Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu

287 1 5 10

290 (2) INFORMATION FOR SEQ ID NO: 4:

RAW SEQUENCE LISTING

DATE: 10/19/2001

PATENT APPLICATION: US/09/921,650A

TIME: 12:07:21

Input Set : N:\Crif3\RULE60\09921650A.txt

Output Set: N:\CRF3\10192001\I921650A.raw

```
292      (i) SEQUENCE CHARACTERISTICS:
293          (A) LENGTH: 11 amino acids
294          (B) TYPE: amino acid
295          (D) TOPOLOGY: linear
297      (ii) MOLECULE TYPE: peptide
299      (v) FRAGMENT TYPE: internal
302      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
304      Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
305      1             5             10
308 (2) INFORMATION FOR SEQ ID NO: 5:
310      (i) SEQUENCE CHARACTERISTICS:
311          (A) LENGTH: 7 amino acids
312          (B) TYPE: amino acid
313          (D) TOPOLOGY: linear
315      (ii) MOLECULE TYPE: peptide
317      (v) FRAGMENT TYPE: internal
320      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
322      Met Pro Lys Arg Pro Arg Pro
323      1             5
326 (2) INFORMATION FOR SEQ ID NO: 6:
328      (i) SEQUENCE CHARACTERISTICS:
329          (A) LENGTH: 569 base pairs
330          (B) TYPE: nucleic acid
331          (C) STRANDEDNESS: double
332          (D) TOPOLOGY: linear
W--> 334      (ii) MOLECULE TYPE: DNA
337      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
339 GAATTCGGGG CCGCGGAGGC TGGATCGGTC CCGGTGTCTT CTATGGAGGT CAAAACAGCG      60
341 TGGATGGCGT CTCCAGGCGA TCTGACGGTT CACTAAACGA GCTCTGCTTA TATAGGTCGA      120
343 GTTTACCACT CCCATCAGT GATAGAGAAA AGTGAAAGTC GAGTTTACCA CTCCCTATCA      180
345 GTGATAGAGA AAAGTGAAAG TCGAGTTTAC CACTCCCTAT CAGTGATAGA GAAAAGTGAA      240
347 AGTCGAGTTT ACCACTCCCT ACCAGTGATA GAGAAAAGTG AAAGTCGAGT TTACCACTCC      300
349 CTATCAGTGA TAGAGAAAAG TGAAAGTCGA GTTTACCACT CCCTATCAGT GATAGAGAAA      360
351 AGTGAAAGTC GAGTTTACCA CTCCCTATCA GTGATAGAGA AAAGTGAAAG TCGAGCTCGG      420
353 TACCCGGGTC GAGTAGGCGT GTACGGTGGG AGGCCTATAT AAGCAGAGCT CGTTTAGTGA      480
355 ACCGTCAGAT CGCCTGGAGA CGCCATCCAC GCTGTTTGA CCTCCATAGA AGACACCGGG      540
357 ACCGATCCAG CCTCCGCGGC CCCGAATTC      569
360 (2) INFORMATION FOR SEQ ID NO: 7:
362      (i) SEQUENCE CHARACTERISTICS:
363          (A) LENGTH: 520 base pairs
364          (B) TYPE: nucleic acid
365          (C) STRANDEDNESS: double
366          (D) TOPOLOGY: linear
W--> 368      (ii) MOLECULE TYPE: DNA
371      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
373 AGATCTGCAG GGTGCTCGG GTTTCGAGGC CACACGCGTC ACCTTAATAT GCGAAGTGGA      60
375 CCGATCTCG AGTTTACCAC TCCCTATCAG TGATAGAGAA AAGTGAAAGT CGAGTTTACC      120
377 ACTCCCTATC AGTGATAGAG AAAAGTGAAA GTCGAGTTTA CCACTCCCTA TCAGTGATAG      180
379 AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGAT AGAGAAAAGT GAAAGTCGAG      240
```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/921,650A

DATE: 10/19/2001

TIME: 12:07:22

Input Set : N:\Crf3\RULE60\09921650A.txt

Output Set: N:\CRF3\10192001\I921650A.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:32 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:37 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:42 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:47 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:52 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:57 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:62 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:67 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:89 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:280 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3
L:334 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:368 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=7
L:508 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11
L:524 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12
L:540 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=13
L:555 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=14
L:571 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=15
L:1023 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=24
L:1039 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=25
L:1073 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=27
L:1433 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=35
L:1475 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=36
L:1511 M:111 C: (47) String data converted to upper case,
M:111 Repeated in SeqNo=37